Correction of sensor crosstalk error in Exhalyzer D multiple-breath washout device significantly impacts outcomes in children with cystic fibrosis

Florian Wyler† (https://orcid.org/0000-0002-1232-1392)
Marc-Alexander H. Oestreich†,2 (https://orcid.org/0000-0001-9641-3691)
Bettina S. Frauchiger1,3 (https://orcid.org/0000-0002-9519-9328)
Kathryn A. Ramsey#1 (https://orcid.org/0000-0003-4574-6917)
Philipp T. Latzin*#1 (https://orcid.org/0000-0002-5239-1571)

† Authors contributed equally to this work as first author
# Authors contributed equally to this work as last author
* Corresponding author

1 Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland
2 Graduate School for Health Sciences, University of Bern, Switzerland.
3 Graduate School of Cellular and Biomedical Sciences, University of Bern, Switzerland

Author contributions:
F.W., M.O., K.R. and P.L contributed to the design of the study; F.W. and M.O. collected and analyzed the data. M.O., B.F and K.R created and organized the databases for retrospective analysis. F.W. drafted the manuscript, with major contributions from M.O, K.R. and P.L. All authors approved the final version of the manuscript.

Running title: Measurement error in a multiple-breath washout device
Corresponding Author:

Prof. Philipp Latzin, MD, PhD
Inselspital, University Children’s Hospital of Bern
Freiburgstrasse 15, 3010 Bern, Switzerland
E-mail: philipp.latzin@insel.ch
Phone: +41 31 632 9353

Total word count: 393 Intro + 492 Methods + 780 Results + 1826 Discussion = 3491

Online Supplement: This manuscript has an online supplement, accessible here:
DOI: https://doi.org/10.6084/m9.figshare.14604846
URL: https://figshare.com/articles/online_resource/_/14604846

Presentations: Preliminary results were presented at the ERS international congress 2020
ABSTRACT

Rationale: Nitrogen multiple-breath washout is an established technique to assess functional residual capacity and ventilation inhomogeneity in the lung. Accurate measurement of gas concentrations is essential for the appropriate calculation of clinical outcomes.

Objectives: We investigated the accuracy of oxygen and carbon dioxide gas sensor measurements used for the indirect calculation of nitrogen concentration in a commercial multiple-breath washout device (Exhalyzer D, Eco Medics AG, Duernten, Switzerland) and its impact on functional residual capacity and lung clearance index.

Methods: High precision calibration gas mixtures and mass spectrometry were used to evaluate sensor output. We assessed the impact of corrected signal processing on multiple-breath washout outcomes in a dataset of healthy children and children with cystic fibrosis using custom analysis software.

Results: We found inadequate correction for the cross sensitivity of the oxygen and carbon dioxide sensors in the Exhalyzer D device. This results in an overestimation of expired nitrogen concentration, and consequently multiple-breath washout outcomes. Breath-by-breath correction of this error reduced the mean (SD) cumulative expired volume by 19.6 (5.0)%, functional residual capacity by 8.9 (2.2)%, and lung clearance index by 11.9 (4.0)%. It also substantially reduced the level of the tissue nitrogen signal at the end of measurements.

Conclusions: Inadequate correction for cross sensitivity in the oxygen and carbon dioxide gas sensors of the Exhalyzer D device leads to an overestimation of functional residual capacity and lung clearance index. Correction of this error is
possible and could be applied by re-analyzing the measurements in an updated software version.

NEW AND NOTEWORTHY

We investigated the sensor accuracy of a prominent nitrogen multiple-breath washout (N₂MBW) device (Eco Medics AG, Duernten, Switzerland) as a possible cause of lack of comparability between outcomes of different MBW devices and methods. We identified an error in the nitrogen concentration calculation of this device, which results in a 10-15% overestimation of primary outcomes, functional residual capacity and lung clearance index. It also leads to a significant overestimation of nitrogen back-diffusion into the lungs.

INTRODUCTION

Cystic fibrosis (CF) is a chronic, genetic disease characterized by lung function decline and respiratory failure. Newborn screening and early interventions for CF have resulted in the majority children with CF having no overt respiratory symptoms and normal spirometry(1). However, structural lung disease is present early on high-resolution chest CT scans and progresses during childhood(2). Therefore, sensitive functional outcomes are needed to monitor disease progression and assess treatment responses in children with CF.

The multiple breath washout (MBW) technique is more sensitive than spirometry to detect early CF lung disease(3, 4). The lung clearance index (LCI) from MBW correlates with underlying structural lung disease and tracks disease progression in children with CF(1, 5-7). LCI has been endorsed as an endpoint in clinical trials in children and adults with CF in North American and Europe(8-10). In young children
with CF, LCI significantly improved in response to disease modifying therapies and hypertonic saline, while spirometry outcomes did not change(11, 12). Therefore, the LCI is a promising endpoint for clinical management and interventional trials in children with CF, even more in the era of CFTR modulator therapies.

The most common MBW technique used in clinical trials is the nitrogen washout (N₂MBW), using the Exhalyzer D device (Eco Medics AG, Duernten, Switzerland)(12, 13). The subject breathes 100% oxygen to wash out resident nitrogen to 2.5% of the starting concentration. However, this technique relies on indirect calculation of N₂ concentration through the measurement of oxygen (O₂) and carbon dioxide (CO₂), which means that small errors in the O₂ or CO₂ concentration could result in instantaneous and cumulative errors in the N₂ concentration, particularly at the end of test criteria(14). Further, as N₂ is soluble in the blood and tissues, there is the potential for N₂ diffusion from the lung tissue into the alveolar spaces during the washout(15). A combination of these effects has been proposed to be the cause of elevated N₂MBW outcomes as compared to SF₆MBW(16). It is essential to address these concerns in N₂ methodology to ensure the appropriate calculation of clinical outcomes.

In this study we investigated the accuracy of indirect nitrogen measurement using the Exhalyzer D device by i) assessing the sensor accuracy of its O₂ and CO₂ sensors, ii) establishing a correction for any observed sensor error and iii) assessing the effect size of the sensor error on clinical outcomes and tissue nitrogen.
METHODS

Study design

This was an experimental study to assess the sensor accuracy of the Exhalyzer D device. We also performed a retrospective analysis of existing N$_2$MBW data to assess the impact of sensor inaccuracy on MBW outcomes. The Ethics Committee of the Canton of Bern, Switzerland approved the study protocol (PB_2017-02139).

i) Sensor accuracy

To assess sensor accuracy over the wide range of concentrations encountered in a N$_2$MBW measurement, we collected experimental data from gas mixture measurements and mass spectrometry measurements. We compared the measured O$_2$ and CO$_2$ concentrations to the known gas mixture concentrations.

Gas mixtures

Sensor accuracy was assessed over a representative range of concentrations present in N$_2$MBW measurements. Twelve technical gas mixtures (Carbagas AG, Muri bei Bern, Switzerland) were used, each containing different combinations of CO$_2$, O$_2$ and N$_2$ concentrations.

Additionally, a series of mass spectrometry measurements was carried out, where N$_2$ was kept at 2% to mimic the MBW end of test condition, while CO$_2$ and O$_2$ were varied (AMIS 2000 Mass Spectrometer, Innovision ApS, Odense, Denmark). Mass spectrometry data was provided by Eco Medics AG, Duernten, Switzerland.

Sensor characteristics

The Exhalyzer D measures both O$_2$ (X3004 OEM sensor, Oxigraf Inc., Sunnyvale, CA, USA) and CO$_2$ (Capnostat 5, Respironics Inc., Wallingford, CT, USA) using
absorption spectroscopy, a technique where the absorption of light is measured at specific frequencies that are characteristic to each gas(17). The absorption spectra of O₂ and CO₂ are affected by a variety of factors, including pressure, temperature, and the presence of other gases(18, 19). This leads to a sensor cross sensitivity, where the presence of O₂ in the gas mixture can affect the absorption spectrum (and therefore measured concentration) of CO₂ and vice versa.

ii) Correction function

We combined the data of the technical gas mixture and mass spectrometry measurements to construct a correction function for the O₂ and CO₂ sensors. We fitted a 2nd-degree two-parameter polynomial through the error for each sensor, as a function of measured O₂ and CO₂ (Eq. E4 in the online supplement). Fitting was performed using MATLAB 2017b (Mathworks, Natick, Massachusetts, USA). This characterization of the measurement error as a function of measured gas concentrations could then be directly used as a correction function for the analysis of MBW measurements. For each combination of O₂ and CO₂ we added the fitted error for each sensor to the respective measured concentrations. Note that when characterizing the error currently present in the sensors, the existing linear CO₂ crosstalk correction of Spiroware 3.2.1 was still applied(20), but when characterizing the correction function, parameters were chosen to replace and improve the existing crosstalk correction.

iii) Impact on outcomes

We characterized the impact of measurement error on MBW outcomes in 357 measurements from 85 healthy children(21) and 62 children diagnosed with CF(5, 22) (Table E4 in the online supplement). We compared outcomes between those
calculated using standard analysis algorithms (replicating Spiroware 3.3) and corrected algorithms using a custom Python script developed by our group.

RESULTS

i) Sensor accuracy

We found that the Exhalyzer D device has O₂ and CO₂ sensor measurement errors which result in overestimated N₂ concentrations. The N₂ error was non-linear and O₂- and CO₂-dependent, with the highest error occurring at very high O₂ concentrations and increasing CO₂ concentrations (Figure 1). Therefore, at the standard end of test conditions (1/40th starting N₂ concentration: end expiratory N₂ at 2%, CO₂ at 5%, O₂ at 93%), the N₂ concentration was significantly overestimated (Table 1). At the original end of test, the Exhalyzer D measures 2% N₂ when the real N₂ concentration was 1.1%. Correction for the sensor error reduces N₂ concentrations at the end of the washout and the end of test criteria was systematically reached earlier (Figure 2). At the new end of test conditions, the relative error in N₂ concentration following correction was estimated to be 44% (2.88% N₂ standard vs 2% N₂ corrected; Table 1).
Table 1: Specific examples of sensor impact on measurement of N₂ in three conditions of interest. The original end of test corresponds to a gas mixture that would be identified as the end of test in standard processing. The second condition corresponds to the new end of test after sensor correction. The third condition contains no real nitrogen. Standard concentrations denote concentrations measured in standard Spiroware 3.3 processing. Corrected concentrations correspond to concentrations after sensor correction is applied. N₂ error summarizes the absolute (abs) difference between N₂ in standard vs. corrected, as well as the relative (rel) error ((standard-corrected)/corrected). The relative contribution of each sensor in [%] to the total error in N₂ concentration is listed under “Contribution.”

<table>
<thead>
<tr>
<th>Condition</th>
<th>Signal Processing</th>
<th>N₂ Error</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard [N₂] [CO₂] [O₂]</td>
<td>Corrected [N₂] [CO₂] [O₂]</td>
<td>abs  rel</td>
</tr>
<tr>
<td>Original end of test [%]</td>
<td>2.00 5.00 93.0</td>
<td>1.10 5.12 93.8</td>
<td>0.90 82.4</td>
</tr>
<tr>
<td>New end of test [%]</td>
<td>2.88 4.88 92.2</td>
<td>2.00 5.00 93.0</td>
<td>0.88 44.1</td>
</tr>
<tr>
<td>No nitrogen [%]</td>
<td>0.88 4.88 94.2</td>
<td>0.00 5.00 95.0</td>
<td>0.88 -</td>
</tr>
</tbody>
</table>
The majority of the error in N\(_2\) measurement (87\% of the error at the test end, see Table 1) originated from the O\(_2\) sensor of the Exhalyzer D device. The O\(_2\) measurement error was non-linear and dependent on the CO\(_2\) concentration (Figure 3A). This error resulted in underestimation of O\(_2\) concentrations, with a greater underestimation with increasing concentrations of CO\(_2\) in the measured range of 0 – 7.5\%. This in turn lead to an overestimation of calculated concentrations of N\(_2\) as described above.

We also found an error in the CO\(_2\) concentration. The CO\(_2\) sensor output was already corrected by a factor that depends on the concentration of O\(_2\) (Eq. E3 in the online supplement). However, the CO\(_2\) sensor seemed to display a different cross-sensitivity than the standard signal processing takes into account (Figure 3B). The residual error appeared to be primarily dependent on CO\(_2\) concentration and only partially dependent on O\(_2\) concentration. The CO\(_2\) error also lead to an overestimation of N\(_2\), however the impact of the CO\(_2\) error was smaller than the O\(_2\) error, making up 13\% of the total sensor error at the end of test conditions (Table 1).

iii) Effect size of sensor correction

Sensor correction impact on MBW outcomes

We re-analyzed 357 MBW measurements from healthy controls (HC) and children with CF using the sensor correction functions outlined above in a custom software. Application of the O\(_2\) and CO\(_2\) sensor correction functions had a significant impact on all MBW outcomes (Table 1). Following the sensor correction, the mean (SD) cumulative expired volume decreased by 19.6 (5.0) \%, FRC decreased by 8.9 (2.2)
%, and LCI decreased by 11.9 (4.0) %. The reduced CEV is explained by lower end-
expiratory concentrations of N\textsubscript{2}, which lead to an earlier end of test (i.e. the criterion
of reaching 1/40\textsuperscript{th} of the initial N\textsubscript{2} concentration is reached earlier). Decreased FRC is
explained by slightly lower concentrations of N\textsubscript{2} throughout the measurement. The
decrease in CEV was more pronounced than for FRC, and with LCI being the ratio of
those two outcomes (LCI = CEV/FRC), this leads to an overall decrease in LCI. The
change in outcomes following sensor correction could vary greatly for individual
measurements (Table 2). However, outcomes before and after the correction over a
large number of measurements correlate strongly. Linear fits of corrected outcomes
vs standard outcomes have R^2 values of 0.997 for FRC, and 0.96 for LCI (Figure 4),
respectively.

The significance of differences in LCI and FRC [L] observed between healthy
controls and children with CF present in the uncorrected data were also present
following sensor correction (Table 2). The change in outcomes following correction
was dependent on the magnitude of the outcomes themselves for both FRC and LCI
(Figure 5, and OLS Figure 1).
### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Corrected</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>mean</td>
<td>SD</td>
</tr>
<tr>
<td><strong>LCI [TO]</strong></td>
<td>All</td>
<td>147</td>
<td>8.33</td>
</tr>
<tr>
<td></td>
<td>HC</td>
<td>85</td>
<td>7.12</td>
</tr>
<tr>
<td></td>
<td>CF</td>
<td>62</td>
<td>9.99</td>
</tr>
<tr>
<td>Difference</td>
<td>-2.87</td>
<td>&lt;0.001</td>
<td>-2.38</td>
</tr>
<tr>
<td><strong>FRC [L]</strong></td>
<td>All</td>
<td>147</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td>HC</td>
<td>85</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>CF</td>
<td>62</td>
<td>1.31</td>
</tr>
<tr>
<td>Difference</td>
<td>0.56</td>
<td>&lt;0.001</td>
<td>0.56</td>
</tr>
<tr>
<td><strong>CEV [L]</strong></td>
<td>All</td>
<td>147</td>
<td>14.9</td>
</tr>
<tr>
<td></td>
<td>HC</td>
<td>85</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>CF</td>
<td>62</td>
<td>15.2</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.58</td>
<td>0.6532</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Table 2: Summary of the differences in Lung Clearance Index (LCI), functional residual capacity (FRC) and cumulative expired volume (CEV) between healthy controls (HC) and patients with cystic fibrosis (CF) in the retrospective dataset before (standard) and after (corrected) the application of the sensor correction function. *unpaired t test; †paired t test. Bold print indicates statistical significance.
Sensor correction impact on tissue nitrogen

We also observed a substantial impact of the sensor corrections on tissue nitrogen. Towards the end of a MBW measurement, the concentration of N\textsubscript{2} in the lung drops so low that diffusion of N\textsubscript{2} from the body becomes a potential concern for the accuracy of the MBW outcomes. In the Exhalyzer D, N\textsubscript{2} concentration is currently overestimated in the presence of CO\textsubscript{2} (i.e. during expirations), even in the complete absence of N\textsubscript{2} (Figure 6B). In conditions reflecting expirations where there is no N\textsubscript{2} exhaled (CO\textsubscript{2} around 5%, rest O\textsubscript{2}), the Exhalyzer D still measures a concentration of N\textsubscript{2} of 0.88\% (Table 1, and Figure 6B, intersection of 5\% line with x-axis). As correction of the O\textsubscript{2} and CO\textsubscript{2} error significantly reduces the N\textsubscript{2} concentration, a significant part of the tissue nitrogen signal at a diffusion equilibrium in long MBW measurements disappears after correction (Figure 6A). The higher the end-expiratory concentrations of CO\textsubscript{2}, the greater this effect (Figure 6B).

DISCUSSION

Summary

We report a significant measurement error in the Eco Medics Exhalyzer D N\textsubscript{2}MBW device. At high concentrations of O\textsubscript{2}, and natural end-expiratory concentrations of CO\textsubscript{2}, the device’s sensors underestimate O\textsubscript{2} and CO\textsubscript{2} gas concentrations and it therefore overestimates end-expiratory concentrations of N\textsubscript{2}. Artificial elevation of N\textsubscript{2} during the washout influences the end of test criterion and causes overestimation of FRC and LCI. It also results in a significant overestimation of measured tissue nitrogen at the end of the test.
i) Sensor accuracy

We are the first group to formally characterize this sensor cross sensitivity error. Previous studies have reported high expiratory N₂ concentrations at the end of long MBW measurements (23-25). While this has predominantly been attributed to the release of N₂ from the lung tissue (15), others have argued that tissue nitrogen alone may not be sufficient to explain the observed concentrations of N₂, and that there may be an additional “offset error” present, speculated to be caused by CO₂-crosstalk with the O₂-sensor (16). We confirm this impact of sensor crosstalk and comprehensively characterize a significant sensor error in the Exhalyzer D device which is primarily responsible for elevated expiratory N₂ concentrations in N₂MBW measurements.

Previous validation studies using *in vitro* lung models as well as the internal testing of Eco Medics AG either did not specifically examine the end of test, end-expiratory conditions examined here, or potentially washed CO₂ out of the validation system before the end of test condition could be reached (26). This may have made it difficult to identify the impact of sensor cross sensitivity in the critical end of test phase of MBW. It is worth noting that individual sensor errors were relatively small (~1% relative error of a sensor reading), even in the most extreme case (low N₂, high CO₂).

However, the indirect calculation of N₂ by the Exhalyzer D device is vulnerable to errors in the high O₂ and high CO₂ concentrations that occur at the end of the MBW measurement (14), leading to a relative N₂ error in this condition of 44%. This measurement error exceeds the recommendations for manufacturers outlined in the ATS/ERS consensus statement of measuring tracer gas concentration within 5% accuracy (27).
The measurements performed here highlight the need for more robust methods of validation for MBW devices. Ideally, such a validation would involve an in vitro lung model that can realistically reproduce the signal dynamics introduced by breathing, and allow for direct comparison of outcomes between devices. In the absence of such a validation system, individual components of MBW devices such as the measurement of tracer gas should be further validated, covering the entire range of concentrations encountered in a MBW measurement. Testing MBW equipment with a wider range of technical gas mixtures with a special emphasis on the test end criterion is highly feasible, and should be a minimum requirement for equipment validation. The specific error in tracer gas measurement described here has been shown to be highly relevant to the Exhalyzer D, but any device that relies on indirect assessment of tracer gas concentrations is potentially vulnerable to cumulative errors in their individual gas sensors.

ii) Correction function

The sensor error observed in this study appears systematic and reproducible across Exhalyzer D devices. The correction function required to correct for the sensor error is simple and has now been implemented in the signal processing of Spiroware (3.3.1), and can also be applied retrospectively to existing data. Sandvik et al. accessed the correction factors and the equations from Eco Medics AG, which have also been published by us in preprint form(28). They applied the equations in a custom-made software version to measurements of healthy infants and toddlers(20). They were able to show that after application of these equations, agreement between $\text{N}_2$-MBW and $\text{SF}_6$-MBW outcomes was closer than without the correction. This and our work suggest that the same equations can be applied to infants and school age children. Notably the correction function suggested here would replace the currently
existing CO$_2$ sensor crosstalk correction. The chosen degree of the polynomial fit constitutes an empirical correction and is a compromise between improving the currently either missing or linear empirical correction and the limits of precision imposed by the intrinsic uncertainty of the reference gas mixtures. To ensure accurate re-analysis of outcomes, this correction function will need to be applied to raw signals on a breath-by-breath basis.

**iii) Effect size of sensor correction**

Sensor correction impact on MBW outcomes

The sensor error described here leads to substantially inflated MBW outcomes. This result provides a new perspective on previously described differences between N$_2$ and SF$_6$ MBW measurements(23-25). It also offers a potential explanation for the differences observed between N$_2$MBW outcomes measured using the Exhalyzer D and devices by other manufacturers such as ndd Medizintechnik AG (Zürich, Switzerland)(29). The primary N$_2$MBW outcomes from the Exhalyzer D were consistently higher than SF$_6$MBW outcomes and N$_2$MBW outcomes from the ndd device. These observations may be partly explained by the systematic overestimation of N$_2$ by the Exhalyzer D reported in this study. The direction of the change after correction suggests that differences between devices will now be smaller. The sensor correction described here has since been used by Sandvik et al. to confirm that agreement between N$_2$MBW and SF$_6$MBW improves upon correction in infants and toddlers(20). In order to validate this in detail, original data need to be reloaded using the sensor correction described here. Fortunately, the N$_2$ error has been an overestimation rather than an underestimation, as measurements can now be re-analyzed without the worry that the trials might not have recorded data long enough to reach the end of test in the corrected measurement.
Notably, using the sensor correction detailed here will lead to substantially shortened measurement times for N\textsubscript{2}MBW. The observed 19.6\% reduction in patient breathing required implies that after the sensor correction, the washout portion of the N\textsubscript{2}MBW measurements would on average be shortened by almost 1/5\textsuperscript{th}. Despite the strong feasibility of N\textsubscript{2}MBW in young children within research studies, challenges have been reported when translating to time-limited busier clinical environments\citep{30}. Shorter test duration may improve this. Previous studies have tried to reduce MBW test duration by using an earlier LCI cut-off\citep{31} (LCI 5\%) or reducing the number of trials used for outcome reporting\citep{32}. The sensor correction described here shortens the N\textsubscript{2}MBW test length without the need to adjusted test protocols or outcomes. With the correction integrated into the Ecomedics software, the reduced test duration for prospective data collection may therefore help to facilitate the transition of N\textsubscript{2}MBW into the clinical setting\citep{5,33}. The effect of the correction function on other MBW indices such as those calculated by concentraton normalised phase III slope analysis (SnIII) remains yet unclear and needs to be examined in future studies.

A major concern that arises with the publication of this study is that it calls into question previously published results obtained using the Exhalyzer D. As the change in outcomes depends on the breathing pattern and CO\textsubscript{2} concentrations, it is difficult for users to predict how much outcomes of a single measurement will change. In addition, the change in outcomes will be higher in children with lung disease and elevated MBW outcomes, compared with healthy children. It is to be expected that effect sizes and confidence intervals of MBW outcomes in such studies will change. This also means that reference values or upper limits of normality generated using the Exhalyzer D device will change\citep{21}. Previously collected results from ongoing
studies will need to be recalculated in order for them to be interpretable alongside
values obtained using this correction.

However, while the impact of the sensor error has effects which are difficult to predict
on the level of individual measurements, the impact on MBW outcomes on a large
enough number of files appears more systematic. Whether or not results from
previous studies are affected can only reliably be elucidated by re-analysis of raw
data. Notably, the impact of the error on outcomes is dependent on their magnitude,
therefore the correction is likely to influence outcomes from individuals with lung
disease more than healthy controls. Even during the retrospective re-analysis within
this study, we observed a change in significance in FRC differences (when
normalized by body weight) between healthy children and children with CF (Table E5
in online supplement). In addition, overestimated values of LCI may have influenced
individual eligibility to enter clinical trials. Re-analysis of MBW measurements used in
clinical trials where drug approval was or is based on affected N₂MBW data should
be prioritized.

Sensor correction impact on tissue nitrogen

It has been hypothesized that towards the end of a N₂MBW test the concentration of
N₂ in the lungs drops so low that a noticeable amount of N₂ diffuses from the body
into the lungs(25, 34). Recent lung modelling work suggests that this N₂ diffusion is
related to local ventilation/perfusion mismatch(35). The results of this current study
suggest that the impact of tissue N₂ diffusion is significantly lower than previously
estimated. Even if no N₂ diffused into the lungs, the Exhalyzer D would still measure
end-expiratory (CO₂ around 5%) concentrations of N₂ of about 0.88%, which would
significantly perturb estimates of tissue nitrogen. The sensor correction functions
introduced in this paper would therefore reduce a substantial part of the observed tissue nitrogen in measurements performed with the Exhalyzer D.

**Strengths and limitations**

Through detailed understanding of the underlying signal processing of the Exhalyzer D we were able to characterize the precise impact of an observed error in gas sensors on the clinical outcomes LCI and FRC. The findings from the technical gases were confirmed by measurements using a mass spectrometer. Using these data, we were able to estimate the impact of the measurement error and develop an appropriate correction function.

The main limitation of this study is the fact that we only had a finite number of gas samples with finite precisions to test the sensors. We chose a selection of gas concentrations from our range of interest which would exhibit cross-sensitivity effects but could ultimately not cover the entire range of concentration combinations in MBW measurements using technical gases. However, the phase of the measurement where sensor accuracy is the most relevant for accurate MBW outcomes is the end of test, whereby the mass spectrometry measurements allowed us to describe the sensor error with high certainty.

**Outlook**

In the process of conducting the research for this paper, we contacted the manufacturer for information regarding their sensor configurations and questions regarding sensor settings and signal processing. The have incorporated the correction described in this manuscript into the signal processing of the new software version of Spiroware 3.3.1, which has since been released.
Conclusion

An error in the cross sensitivity correction between the oxygen and carbon dioxide gas sensors of the Exhalyzer D device leads to an overestimation of FRC and LCI. Correction of this error is possible but needs to be applied breath-by-breath by re-analyzing the measurements in an updated version of the Spiroware analysis software.
ACKNOWLEDGMENTS

The authors thank Daniel Oberli and Ruedi Isler (Eco Medics AG, Switzerland) for confirming the sensor error after we detected and informed them about it. We also thank Daniel Oberli and Ruedi Isler for working with us to finalize the correction equation based on the technical gas measurements performed by us and the mass spectrometry measurements performed by them, which have now been implemented into the commercial software. We thank all our patients and study participants as well as their families for allowing their MBW data to be used for research. In addition, we thank the lab technicians Sandra Lüscher, Bettina Vessaz, Sharon Krattinger, and Gisela Wirz for their contribution by performing MBW measurements. Special thanks go to the physicians from our department.

GRANTS, GIFTS, EQUIPMENT, DRUGS

Eco Medics AG (Duernten, Switzerland) provided a research version of their commercial software Spiroware 3.3, provided information on signal processing algorithms, and helped with the acquisition of mass spectrometry measurements. This project was funded by the Swiss National Science Foundation, Grant Nr. 182719 (P. Latzin) and 168173 (K. Ramsey).
REFERENCES


20. Sandvik RM, Gustafsson PM, Lindblad A, Robinson PD, and Nielsen KG. Improved agreement between N(2) and SF(6) multiple breath washout in healthy infants and toddlers with improved EXHALYZER D(®) sensor performance. *Journal of applied physiology (Bethesda, Md : 1985)* 2021.


Figure 1: N₂ error as a function of N₂ in the gas mixtures. The N₂ error here is the absolute difference between measured N₂ and reference N₂ as a function of reference N₂ in the gas mixtures. Dashed curves represent the combined fits through the errors of the individual gas sensor errors, for selected concentrations of CO₂. Dotted vertical line indicates the end of test condition. The color shading indicates the reference CO₂ concentrations of the gas mixtures. Dots represent mean of 6 measurements (triplicates on 2 devices) performed with 12 technical gas mixtures as reference (CO₂: 0%, 2.5%, 7.5%, O₂: 30%, 60%, 90%, Rest, N₂: Rest), triangles represent mean of 3 mass spectrometry reference measurements of 7 mixtures at the end of test condition (N₂: 2%, CO₂: 0%, 1%, 2%, 3%, 4%, 5%, 6%, O₂: Rest). Error bars represent SD of measurements for each mixture. For an overview of the gas mixtures see OLS (Technical gases and Table 2).
**Figure 2**: Illustration of the effect of the sensor correction on the N\textsubscript{2} signal and consequently on the end of test in an example MBW measurement. Traced in gray is the signal output of the standard signal processing, the corrected signal is shown in black. Vertical dashed lines represent the end of test for the original standard and corrected measurement respectively. The dashed horizontal line corresponds to 1/40\textsuperscript{th} of the initial N\textsubscript{2} concentration (end of test, ca. 2\% N\textsubscript{2}). Dashed vertical lines represent the original end of test (end of test condition reached in standard processing) and new end of test (end of test condition reached in corrected processing). (A) Time course of N\textsubscript{2} throughout a standard MBW measurement. (B) Zoom into the critical period of end of test determination. In this example the test ends 5 breaths earlier in the corrected measurement compared to standard.
Figure 3: Observed absolute error between reference and measured gas concentrations. (dots: mean of error of one gas mixture, error bars: +/- SD of error). Curves represent a two parameter quadratic polynomial fitted through the error values (see OLS for details), represented here as dashed curves for given CO₂ concentrations. Dots represent mean of 6 measurements (triplicates on 2 devices) performed with 12 technical gas mixtures as reference (CO₂: 0%, 2.5%, 7.5%, O₂: 30%, 60%, 90%, Rest, N₂: Rest), triangles represent mean of 3 mass spectrometry reference measurements of 7 mixtures at the end of test condition (N₂: 2%, CO₂: 0%, 1%, 2%, 3%, 4%, 5%, 6%, O₂: Rest). Error bars represent SD of measurements for each mixture. For an overview of the gas mixtures see OLS (Technical gases and Table 2). (A) Absolute O₂ error as a function of O₂ and CO₂ concentration, (B) Absolute CO₂ error as a function of O₂ and CO₂ concentration.
Figure 4: Multiple-breath washout outcomes (A) Lung Clearance Index (LCI) and (B) functional residual capacity (FRC) after sensor correction (corrected) vs standard (standard; Spiroware 3.3) in healthy controls (HC) and patients with cystic fibrosis (CF). Solid black line indicates line of equality, dashed line represents a linear fit through the data points.
Figure 5: Bland-Altman plot of the absolute difference (corrected – standard) of multiple-breath washout outcomes (A) Lung Clearance Index (LCI) in turnover [TO] and (B) functional residual capacity (FRC) in liter [L] of healthy controls (HC) and patients with cystic fibrosis (CF) due to sensor correction, plotted against the mean outcomes (mean of corrected and standard).
Figure 6: Illustration of the effect of the sensor correction function on nitrogen measurement in the late phase of MBW tests. (A) Example of the equilibrium N₂ reached in a very long continued MBW measurement, displaying a greatly decreased N₂-back-diffusion equilibrium (tissue nitrogen). (B) Corrected N₂ plotted against standard N₂ in conditions around the end of test condition (2% N₂).
A

[Graph A]

- Standard
- Corrected
- Test-end criterion

Time [s]

N₂ [%]

B

[Graph B]

- CO₂ = 0%
- CO₂ = 2.5%
- CO₂ = 5%
- CO₂ = 7.5%
- N₂ = 2%

N₂ corrected [%]

N₂ standard [%]
Table 1: Specific examples of sensor impact on measurement of N₂ in three conditions of interest. The original end of test corresponds to a gas mixture that would be identified as the end of test in standard processing. The second condition corresponds to the new end of test after sensor correction. The third condition contains no real nitrogen. Standard concentrations denote concentrations measured in standard Spiroware 3.3 processing. Corrected concentrations correspond to concentrations after sensor correction is applied. N₂ error summarizes the absolute (abs) difference between N₂ in standard vs. corrected, as well as the relative (rel) error ((standard-corrected)/corrected). The relative contribution of each sensor in [%] to the total error in N₂ concentration is listed under “Contribution.”

<table>
<thead>
<tr>
<th>Condition</th>
<th>Signal Processing</th>
<th>N₂ Error</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[N₂] [CO₂] [O₂]</td>
<td>[N₂] [CO₂] [O₂]</td>
<td>abs rel [CO₂] [O₂]</td>
</tr>
<tr>
<td>Original end of test [%]</td>
<td>2.00 5.00 93.0</td>
<td>1.10 5.12 93.8</td>
<td>0.90 82.4 13.5 86.5</td>
</tr>
<tr>
<td>New end of test [%]</td>
<td>2.88 4.88 92.2</td>
<td>2.00 5.00 93.0</td>
<td>0.88 44.1 13.3 86.7</td>
</tr>
<tr>
<td>No nitrogen [%]</td>
<td>0.88 4.88 94.2</td>
<td>0.00 5.00 95.0</td>
<td>0.88 - - -</td>
</tr>
</tbody>
</table>

| Condition                  |  |  |  |  |
Table 2

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>p-value*</th>
<th>Mean</th>
<th>SD</th>
<th>p-value*</th>
<th>Mean rel [%]</th>
<th>95% CI [%]</th>
<th>p-value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCI [TO]</td>
<td>All</td>
<td>147</td>
<td>8.33</td>
<td>2.05</td>
<td>7.31</td>
<td>1.7</td>
<td></td>
<td>1.02</td>
<td>11.9</td>
<td>11.2 - 12.5</td>
</tr>
<tr>
<td></td>
<td>HC</td>
<td>85</td>
<td>7.12</td>
<td>0.51</td>
<td>6.30</td>
<td>0.4</td>
<td></td>
<td>0.82</td>
<td>11.3</td>
<td>10.6 - 12.1</td>
</tr>
<tr>
<td></td>
<td>CF</td>
<td>62</td>
<td>9.99</td>
<td>2.21</td>
<td>8.69</td>
<td>1.8</td>
<td></td>
<td>1.30</td>
<td>12.6</td>
<td>11.4 - 13.8</td>
</tr>
<tr>
<td>Difference</td>
<td>-2.87</td>
<td>&lt;0.001</td>
<td>-2.38</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRC [L]</td>
<td>All</td>
<td>147</td>
<td>1.63</td>
<td>0.87</td>
<td>1.49</td>
<td>0.80</td>
<td></td>
<td>0.14</td>
<td>8.9</td>
<td>8.6 - 9.3</td>
</tr>
<tr>
<td></td>
<td>HC</td>
<td>85</td>
<td>1.87</td>
<td>0.95</td>
<td>1.73</td>
<td>0.89</td>
<td></td>
<td>0.14</td>
<td>7.9</td>
<td>7.6 - 8.1</td>
</tr>
<tr>
<td></td>
<td>CF</td>
<td>62</td>
<td>1.31</td>
<td>0.61</td>
<td>1.17</td>
<td>0.53</td>
<td></td>
<td>0.14</td>
<td>10.4</td>
<td>9.7 - 11.0</td>
</tr>
<tr>
<td>Difference</td>
<td>0.56</td>
<td>&lt;0.001</td>
<td>0.56</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEV [L]</td>
<td>All</td>
<td>147</td>
<td>14.9</td>
<td>7.4</td>
<td>11.9</td>
<td>5.7</td>
<td></td>
<td>3.03</td>
<td>19.6</td>
<td>18.8 - 20.4</td>
</tr>
<tr>
<td></td>
<td>HC</td>
<td>85</td>
<td>14.6</td>
<td>6.7</td>
<td>12.0</td>
<td>5.6</td>
<td></td>
<td>2.67</td>
<td>18.2</td>
<td>17.4 - 19.0</td>
</tr>
<tr>
<td></td>
<td>CF</td>
<td>62</td>
<td>15.2</td>
<td>8.3</td>
<td>11.7</td>
<td>5.8</td>
<td></td>
<td>3.54</td>
<td>21.5</td>
<td>20.1 - 23.0</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.58</td>
<td>0.6532</td>
<td>0.29</td>
<td>0.7601</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Summary of the differences in Lung Clearance Index (LCI), functional residual capacity (FRC) and cumulative expired volume (CEV) between healthy controls (HC) and patients with cystic fibrosis (CF) in the retrospective dataset before (standard) and after (corrected) the application of the sensor correction function. *unpaired t test; †paired t test. Bold print indicates statistical significance.